

WHAT IS CLAIMED IS:

- 1 1. A method for finding one or more target biometric samples that are similar to
2 or match a query biometric sample, comprising:
 - 3 generating a query feature vector from a query biometric vector that
4 represents said query biometric sample, said query feature vector comprising a
5 plurality of features and said query biometric vector comprising a set of
6 characteristics; and
 - 7 comparing said query feature vector to a plurality of target feature vectors,
8 each target feature vector representing a respective target biometric sample;
9 wherein,
 - 10 a target biometric sample is a potential match to said query biometric sample
11 when a threshold number of features in the target feature vector that corresponds to
12 said target biometric sample are identical to features in said query biometric vector.
- 1 2. The method of claim 1, wherein the generating comprises:
 - 2 extracting a set of tiles that represents said query biometric sample, each tile
3 including a plurality of characteristics;
 - 4 assigning an identification element to each said tile in said set of tiles;
 - 5 selecting a predetermined number of the identification elements;
 - 6 partitioning the selected identification elements into a plurality of partitioned
7 groups; and
 - 8 applying a fingerprint function to each partitioned group to generate a set of
9 features, said set of features representing said query feature vector.
- 1 3. The method of claim 2, wherein said extracting step further comprises:
 - 2 canonicalizing each characteristic associated with said query biometric
3 sample to create a plurality of canonicalized characteristics; and
 - 4 obtaining a pseudo randomly selected subset of said plurality of
5 canonicalized characteristics to form each said plurality of characteristics included in
6 said set of tiles.

1 4. The method of claim 3, wherein a characteristic associated with said query
2 biometric sample is a variable characteristic and said canonicalizing step includes
3 expansion of said variable characteristic using a predetermined function.

1 5. The method of claim 3, wherein said canonicalization of each said
2 characteristic associated with said query biometric sample includes application of a
3 respective weight to each said characteristic, the respective weight determining a
4 number of copies of said characteristic that are present in said plurality of
5 canonicalized characteristics.

1 6. The method of claim 1, wherein each target feature vector in said plurality of
2 target feature vectors is generated by:

3 extracting a set of tiles that represent a respective target biometric sample,
4 each tile including a plurality of characteristics;
5 assigning an identification element to each said tile in said set of tiles;
6 selecting a predetermined number of the identification elements;
7 partitioning the selected identification elements into a plurality of groups; and
8 applying a fingerprint function to each group to generate a set of features,
9 said set of features representing said target feature vector.

1 7. The method of claim 6, wherein said extracting step further comprises:
2 canonicalizing the characteristics associated with said respective biometric
3 sample to create a plurality of canonicalized characteristics; and
4 obtaining a pseudo randomly selected subset of said plurality of
5 canonicalized characteristics to form each said plurality of characteristics included in
6 said set of tiles.

1 8. The method of claim 7, wherein a characteristic associated with said
2 respective biometric sample is a variable characteristic and said canonicalizing step
3 includes expansion of said variable characteristic using a predetermined function.

1 9. The method of claim 7, wherein said canonicalization of each said
2 characteristic associated with said respective biometric sample includes application

3 of a respective weight to each said characteristic, the respective weight determining
4 a number of copies of said characteristic that are present in said plurality of
5 canonicalized characteristics.

1 10. The method of claim 2, wherein the selecting includes selecting the
2 identification elements using a selection function.

1 11. The method of claim 6, wherein the selecting includes selecting the
2 identification elements using a selection function.

1 12. The method of claim 2, wherein each of the set of features is a one-way hash
2 of a group.

1 13. The method of claim 6, wherein each of the set of features is a one-way hash
2 of a group.

1 14. The method of claim 1, wherein the comparing includes:
2 extracting from a database a set of target feature vectors, each extracted
3 target feature vector including at least one feature that is in common with a feature
4 in said query feature vector;

5 locating each target feature vector in the set of target feature vectors in a
6 count hash table, the count hash table including target feature vector identifications
7 corresponding to target feature vectors that share at least one feature in common
8 with the target feature vector, the count hash table including a count of matching
9 features for each target feature vector represented by the table; and

10 incrementing the count of matching features for each located target feature
11 identification.

1 15. A computer program product for use in conjunction with a computer system,
2 the computer program product comprising a computer readable medium and a
3 computer program mechanism embedded therein, the computer program
4 mechanism comprising:

5 a data structure including a plurality of target feature vectors, each target
6 feature vector representing a respective target biometric sample;
7 a module for finding one or more said respective target biometric samples
8 that are similar to or match a query biometric sample, said module including
9 instructions for:

10 generating a query feature vector from a query biometric vector that
11 represents said query biometric sample, said query feature vector comprising a
12 plurality of features and said query biometric vector comprising a set of
13 characteristics; and

14 comparing said query feature vector to said plurality of target feature vectors;
15 wherein,

16 a target biometric sample is a potential match to said query biometric sample
17 when a threshold number of features in the target feature vector that corresponds to
18 said target biometric sample are identical to features in said query biometric vector.

1 16. The computer program product of claim 15, wherein the instructions for
2 generating comprise:

3 instructions for extracting a set of tiles that represents said query biometric
4 sample, each tile including a plurality of characteristics;

5 instructions for assigning an identification element to each said tile in said set
6 of tiles;

7 instructions for selecting a predetermined number of the identification
8 elements;

9 instructions for partitioning the selected identification elements into a plurality
10 of partitioned groups; and

11 instructions for applying a fingerprint function to each partitioned group to
12 generate a set of features, said set of features representing said query feature
13 vector.

1 17. The computer program product of claim 20, wherein said instructions for
2 extracting further comprise:

3 instructions for canonicalizing each characteristic associated with said query
4 biometric sample to create a plurality of canonicalized characteristics; and

5 instructions for obtaining a pseudo randomly selected subset of said plurality
6 of canonicalized characteristics to form each said plurality of characteristics included
7 in said set of tiles.

1 18. The computer program product of claim 17, wherein a characteristic
2 associated with said query biometric sample is a variable characteristic and said
3 instructions for canonicalizing includes expansion of said variable characteristic
4 using a predetermined function.

1 19. The computer program product of claim 17, wherein said instructions for
2 canonicalization of each said characteristic associated with said respective biometric
3 sample includes application of a respective weight to each said characteristic, the
4 respective weight determining a number of copies of said characteristic that are
5 present in said plurality of canonicalized characteristics.

1 20. The computer program product of claim 15, wherein said module further
2 includes instructions for creating each target feature vector in said plurality of target
3 feature vectors, said instructions for creating including:

4 instructions for extracting a set of tiles that represent a respective target
5 biometric sample, each tile including a plurality of characteristics;

6 instructions for assigning an identification element to each said tile in said set
7 of tiles;

8 instructions for selecting a predetermined number of the identification
9 elements;

10 instructions for partitioning the selected identification elements into a plurality
11 of groups; and

12 instructions for applying a fingerprint function to each group to generate a set
13 of features, said set of features representing said target feature vector.

1 21. The computer program product of claim 20, wherein said instructions for
2 extracting further comprise:

3 instructions for canonicalizing the characteristics associated with said
4 respective biometric sample to create a plurality of canonicalized characteristics; and

5 instructions for obtaining a pseudo randomly selected subset of said plurality
6 of canonicalized characteristics to form each said plurality of characteristics included
7 in said set of tiles.

1 22. The computer program product of claim 21, wherein a characteristic
2 associated with said respective biometric sample is a variable characteristic and said
3 instructions for canonicalizing include expansion of said variable characteristic using
4 a predetermined function.

1 23. The computer program product of claim 21, wherein said instructions for
2 canonicalization of each said characteristic associated with said respective biometric
3 sample includes application of a respective weight to each said characteristic, the
4 respective weight determining a number of copies of said characteristic that are
5 present in said plurality of canonicalized characteristics.

1 24. The computer program product of claim 16, wherein said instructions for
2 selecting include instructions for selecting the identification elements using a
3 selection function.

1 25. The computer program product of claim 20, wherein said instructions for
2 selecting include instructions for selecting the identification elements using a
3 selection function.

1 26. The computer program product of claim 16, wherein each of the set of
2 features is a one-way hash of a group.

1 27. The computer program product of claim 20, wherein each of the set of
2 features is a one-way hash of a group.

1 28. The computer program product of claim 15, wherein the instructions for
2 comparing include:

3 instructions for extracting from said data structure a set of target feature
4 vectors, each extracted target feature vector including at least one feature that is in
5 common with a feature in said query feature vector;
6 instructions for locating each target feature vector in the set of target feature
7 vectors in a count hash table, the count hash table including target feature vector
8 identifications corresponding to target feature vectors that share at least one feature
9 in common with the target feature vector, the count hash table including a count of
10 matching features for each target feature vector represented by the table; and
11 instructions for incrementing the count of matching features for each located
12 target feature identification.

1 29. A system for identifying potential matches between a query biometric sample
2 and a plurality of target biometric samples comprising:
3 a memory storing a plurality of target feature vectors, each target feature
4 vector representing a respective target biometric sample in said plurality of target
5 biometric samples;
6 a central processing unit for executing a software module, the software
7 module including instructions for:
8 generating a query feature vector from a query biometric vector, said query
9 biometric vector representing said query biometric sample, said query feature vector
10 comprising a plurality of features and said query biometric vector comprising a set of
11 characteristics; and
12 comparing said query feature vector to said plurality of target feature vectors;
13 wherein,
14 a target biometric sample is a potential match to said query biometric sample
15 when a threshold number of features in the target feature vector that corresponds to
16 said target biometric sample are identical to features in said query biometric vector.